

# STA257

Neil Montgomery | HTML is official | PDF versions good for in-class use only  
Last edited: 2016-10-24 13:19

## (more) gamma (distribution) trivia

Recall:

$$\Gamma(\alpha) = \int_0^{\infty} u^{\alpha-1} e^{-u} du$$

Also recall density of  $\text{Gamma}(\alpha, \lambda)$  distributions:

$$f(x) = \frac{\lambda^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$$

This would be used to compute probabilities.  $f(x)$  does not have a closed-form anti-derivative (as usual), so a numerical routine is used (as usual).

$\alpha = n/2$  and  $\lambda = 1/2$  is called  $\chi_n^2$  and has applications in statistics. (More today.)

## another fun Poisson process fact

Suppose that at some fixed time  $t$  of a Poisson process we know  $N(t) = 1$ . In other words, exactly one "event" occurred at some time before  $t$ .

This occurrence time is a random variable. Call it  $X$ . What is its distribution?

Let's try to derive its cdf:

$$F(x) = P(X \leq x) = \begin{cases} 0 & : x < 0 \\ ??? & : 0 \leq x \leq t \\ 1 & : x \geq t \end{cases}$$

## exercise

Suppose that at some fixed time  $t > 0$  of a Poisson process, we know  $N(t) = n$ , but we don't know when the  $n$  events occurred.

Fix another time  $s$  with  $0 < s < t$ .

Find the distribution of the number  $X$  of events that occurred inside  $[0, s]$ .

# "Normal" distributions

A (continuous) random variable  $X$  has a "Normal" (or "Gaussian") distribution if its density is:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

Normal distributions are indexed by two parameters: the "mean"  $-\infty < \mu < \infty$  and the "standard deviation"  $\sigma > 0$ .

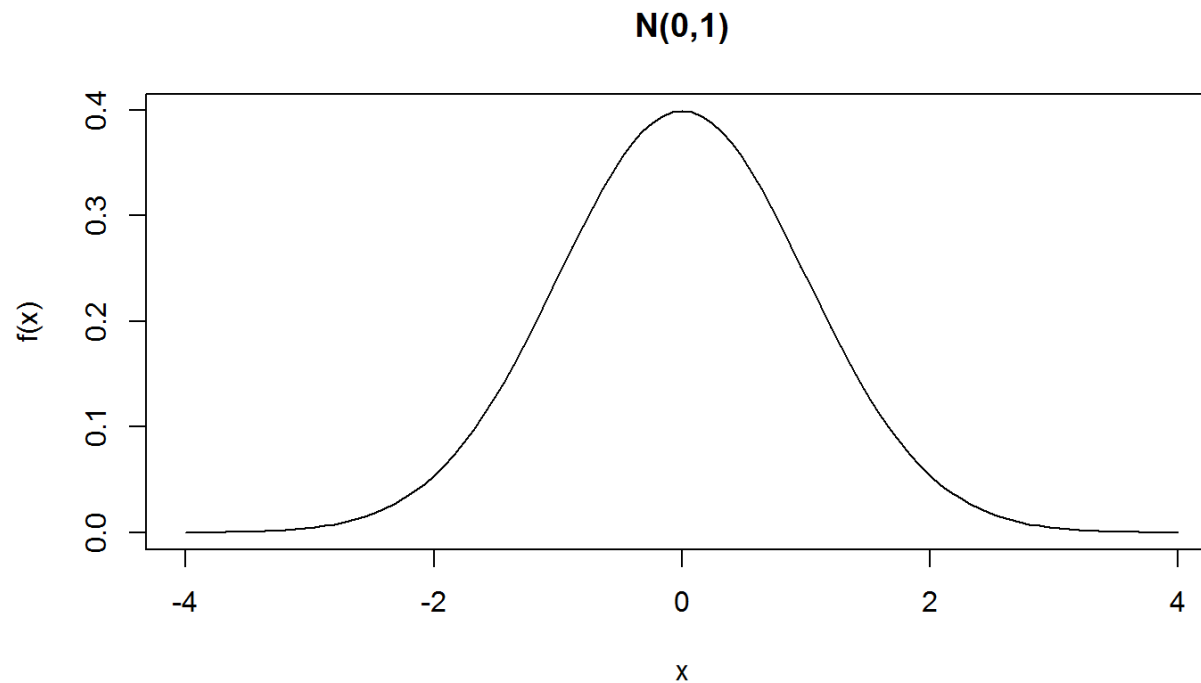
It is also common to index by  $\mu$  and  $\sigma^2 = (\sigma)^2$  called the "variance".

This is called a "reparametrization".

Common notation:  $X \sim N(\mu, \sigma)$  or  $X \sim N(\mu, \sigma^2)$  (I *always* use the latter.)

# free picture of a Normal density

They all have this shape:



# density properties; calculations; applications

Maximum is attained at  $x = \mu$  and inflection points are at  $\mu \pm \sigma$  (basic calculus).

The density is *symmetric* around  $\mu$ , which means  $f(\mu - a) = f(\mu + a)$ .

The Normal densities do not have closed-form anti-derivates (as usual) so a numerical routine is used (as usual).

You'll also have to get used to using tables.

Direct use: modeling things that are the result of "sums" of small independent contributions.

But the indirect use is far more important, and is the reason why they get all the love rather than the Poisson process family. (More later.)

# N(0,1) calculation examples

Suppose  $Z \sim N(0, 1)$ . ("The standard normal"). The book includes a table of  $P(Z \leq z_p)$  for  $z_p \in \{0.00, 0.01, \dots, 3.49\}$  with  $P(Z \leq z_p) = p$ .

From this table, for example:

$$P(Z \leq 1) = 0.8413 = P(Z \geq -1)$$

$$P(-1 < Z < 1) = 0.6827$$

$$P(Z \leq z_p) = 0.91 \iff z_p \approx 1.34$$

and plenty more. Become handy with this table.



# "direct" normal model example

The SAT is a standardized test taken mainly by USA high school students, and some others.

The scores can be modeled using a normal distribution (why?) with, according to the May 2016 report, mean 1020 and standard deviation 194.

What is the probability that a student will have a score that exceeds 1400?

The answer is 0.1179 according to my computer.

But we don't have a  $N(1020, 194)$  table, which leads us to...

distributions of functions of random  
variables (with a focus on  
continuous r.v.)

# functions of random variables: why?

Some of your calculus life was spent pondering things like  $f(g(x))$  and their derivatives and integrals.

Now we'll ponder things like  $g(X)$  focusing, of course, on the resulting *distributions*.

A common simple example would be a linear transformation ("unit change") with  $g(x) = a + bx$  and:

$$Y = g(X) = a + bX$$

Given the distribution of  $X$ , what will be the distribution of  $Y$ ?

"DOFORV" - "distribution of a function of a random variable"

# DOFORV method I - use the cdf

The cdf is one way to characterize a distribution.

For example, suppose we have the cdf of  $X$  given by  $F_x(x) = P(X \leq x)$  and we want the distribution of  $Y = g(X) = a + bX$  for  $b \neq 0$ .

If we can find the cdf of  $Y$ , that gives us what we want:

$$F_y(y) = \begin{cases} F_x\left(\frac{y-a}{b}\right) & : \text{if } b > 0 \\ 1 - F_x\left(\frac{y-a}{b}\right) & : \text{if } b < 0 \end{cases}$$

Then differentiate for density:

$$f_y(y) = f_x\left(\frac{y-a}{b}\right) \cdot \frac{1}{|b|}$$

## example: the normal distributions

Suppose  $X \sim N(\mu, \sigma^2)$  and consider  $g(x) = \frac{x-\mu}{\sigma}$ . Determine the distribution of  $Z = g(X)$ .

We can go straight to the density to see  $Z \sim N(0, 1)$ .

Now we can use the table to finish the SAT example, in which  $X \sim N(1020, 194)$  and we wanted  $P(X > 1250)$ .

$Z \sim N(0, 1)$  is called the "standard normal distribution". Its cdf gets its own special symbol:  $F_z(z) = P(Z \leq z) = \Phi(z)$ .

Apparently its density gets to be called  $\phi(z)$  but this isn't as common.

Exercise: if  $Z$  is standard normal, find the distribution of  $X = \mu + \sigma Z$ .

## example: uniform

Suppose  $X \sim \text{Unif}[0, 1]$ . Determine the distribution of  $Y = a + bX$  with  $b \neq 0$ .

Exercise: Suppose  $X \sim \text{Unif}[a, b]$ . Find  $c$  and  $d$  so that  $Y = c + dX$  is  $\text{Unif}[0, 1]$ .

Exercise: Suppose  $X \sim \text{Exp}(1)$ . Determine the distribution of  $Y_1 = \lambda X$  and  $Y_2 = a + \lambda X$ . Draw pictures of them. Do consider  $\lambda < 0$  (perhaps without any obvious application but still valid mathematically.)

(Note: the last part will result in a distribution we haven't given a name to, which is fine.)

# DOFORV method I - cdf of square of r.v.

Squaring a random variable will turn out to have important applications.

Suppose  $X$  has cdf  $F_x$  and density  $f_x$ , and  $g(x) = x^2$ . Determine the distribution of  $Y = g(X) = X^2$  using the cdf method.

$$F_y(y) = F_x(\sqrt{y}) - F_x(-\sqrt{y})$$

$$f_y(y) = \frac{1}{2\sqrt{y}} (f_x(\sqrt{y}) + f_x(-\sqrt{y}))$$

Example:  $Z \sim N(0, 1)$  and  $Y = Z^2$ . Then  $Y$  has a  $\text{Gamma}(\frac{1}{2}, \frac{1}{2})$  distribution, a.k.a.  $\chi_1^2$ , with a bonus fun fact for free.

## DOFORV method 2 - direct "theorem"

I am delighted that the book downplays this method as not as easy to use. Nor do I recommend it for practical use.

Theorem: Given  $X$  with density  $f_x(x)$  and  $g$  monotone and differentiable with inverse  $g^{-1}$  where  $f_x(x) > 0$ , let  $Y = g(X)$ . Then:

$$f_y(y) = f_x(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|$$

This theorem can be extended to non-monotonic  $g$ .

Proof: ...



# DOFORV - two other proofs

The theorem can be proved in two other ways that are straight outta calculus.

The proofs indeed all look very similar.

Proof 2: uses the "change of variables" method from integration (emphasizes my advice to always think of a density as living in an integral.)

Proof 3: uses the fundamental theorem of calculus.

# a seemingly strange example

The techniques apply to any continuous r.v.  $X$  and to any differentiable, invertible  $g(x)$ .

So let's consider  $X \sim \text{Exp}(\lambda)$  and let  $g(x) = 1 - e^{-\lambda x}$ . It turns out  $Y \sim \text{Unif}[0, 1]$ .

The function  $g$  was not chosen by accident—it is precisely the cdf  $F_X(x)$  of  $X$ .

Theorem: If  $X$  is continuous and has cdf  $F_X(x)$  then  $Y = F_X(X)$  will have a uniform distribution on  $[0, 1]$ .

Proof: ...